

Conference Abstract

Integration of Biodiversity Linked Data and Web APIs using SPARQL Micro-Services

Franck Michel[‡], Catherine Faron-Zucker[‡], Fabien Gandon[‡]

[‡] Université Côte d'Azur, CNRS, Inria, I3S, Sophia-Antipolis, France

Corresponding author: Franck Michel (franck.michel@cnrs.fr)

Received: 03 Apr 2018 | Published: 22 May 2018

Citation: Michel F, Faron-Zucker C, Gandon F (2018) Integration of Biodiversity Linked Data and Web APIs using SPARQL Micro-Services. Biodiversity Information Science and Standards 2: e25481.

<https://doi.org/10.3897/biss.2.25481>

Abstract

Web APIs (Application Programming Interface) are a common means for Web portals and data producers to enable HTTP-based, machine-processable access to their data. They are a prominent source of information*1 pertaining to topics as diverse as scientific information, social networks, entertainment or finance. The methods of Linked Data (Heath and Bizer 2011) similarly aim to publish machine-readable data on the Web, while connecting related resources within and between datasets, thereby creating a large distributed knowledge graph. Today, the biodiversity community is increasingly adopting the Linked Data principles to publish data such as trait banks, museum collections and taxonomic registers (Parr et al. 2016, Baskauf et al. 2016). However, standard approaches are still missing to combine disparate representations coming from both Linked Data interfaces and the manifold Web APIs that were developed during the last two decades to expose legacy biodiversity databases on the Web.

The *SPARQL Micro-Service* architecture (Michel et al. 2018) tackles the goal of reconciling Linked Data interfaces and Web APIs. It proposes a lightweight method to query a Web API using SPARQL (Harris and Seaborne 2013), the Semantic Web standard to query knowledge graphs expressed in the Resource Description Framework (RDF). A SPARQL micro-service provides access to a small RDF graph, typically resource-centric, that it builds at run-time by transforming a fraction of the whole dataset served by the Web API into RDF triples. Furthermore, Web APIs traditionally rely on internal, proprietary resource

identifiers that are unsuited for use as Uniform Resource Identifiers (URIs). To address this concern, a SPARQL micro-service can assign a URI to a Web API resource, allowing an application to look up this URI and get a description of the resource in return (this process is referred to as *dereferencing*).

In this demo, we wish to showcase the value of SPARQL micro-services in the biodiversity domain. We first query TAXREF-LD, a Linked Data representation of the French taxonomic register of living beings (Michel et al. 2017), to retrieve information about a given taxon. Then, we demonstrate how we can enrich our knowledge about this taxon with various types of data retrieved on-the-fly from multiple Web APIs:

- trait data from the Encyclopedia of Life trait bank (Parr et al. 2016),
- articles or books from the [Biodiversity Heritage Library](#),
- audio recordings from the [Macaulay scientific media archive](#),
- photos from the [Flickr photography social network](#), and
- music tunes from [MusicBrainz](#).

Different visualizations are demonstrated, ranging from raw RDF triples to Web pages generated dynamically and integrating heterogeneous data, as suggested in Fig. 1. Depending on the audience’s interests, we shall touch upon the alignment of Web APIs’ proprietary vocabularies with well-adopted thesauri or ontologies, or more technical concerns *e.g.* related to the effort required to deploy a new SPARQL micro-service.

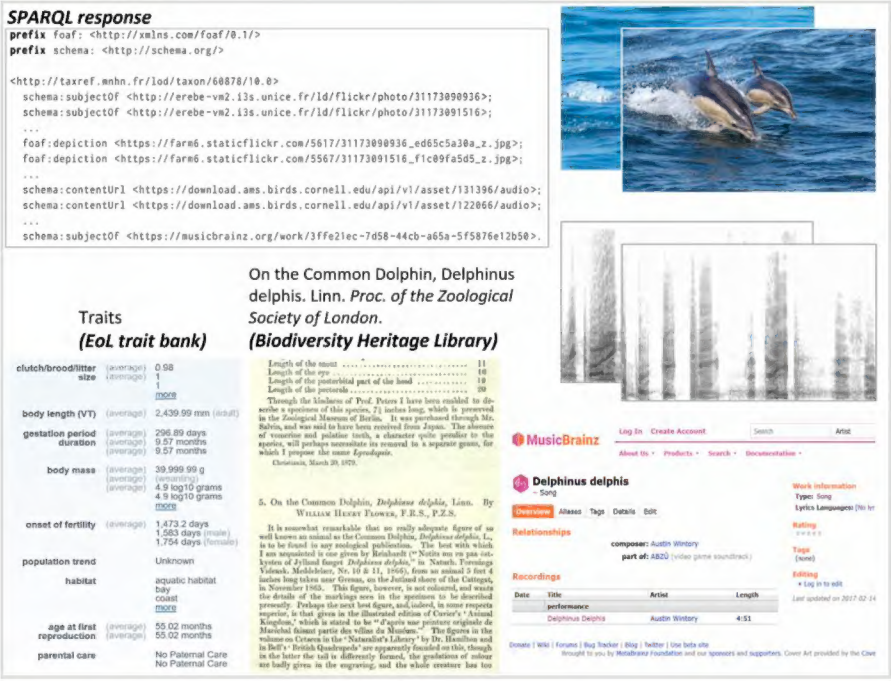


Figure 1. Visualization of multiple types of data retrieved from Web APIs about taxon *Delphinus delphis*.

Keywords

Web API, SPARQL, data integration, linked data, JSON-LD

Presenting author

Franck Michel is a research engineer at the University Côte d'Azur, CNRS, France. His research topics notably concern the integration and federation of heterogeneous data sources using Semantic Web technologies, and their publication in the Web of Data.

References

- Baskauf S, Wiczorek J, Deck J, Webb C (2016) Lessons Learned from Adapting the Darwin Core Vocabulary Standard for Use in RDF. *Semantic Web – Interoperability, Usability, Applicability* 7 (6): 617-627.
- Harris S, Seaborne A (2013) SPARQL 1.1 Query Language. W3C Recommendation.
- Heath T, Bizer C (2011) *Linked Data: Evolving the Web into a Global Data Space*. 1st. Morgan & Claypool
- Michel F, Gargominy O, Tercerie S, Faron-Zucker C (2017) A Model to Represent Nomenclatural and Taxonomic Information as Linked Data. Application to the French Taxonomic Register, TAXREF. *Proceedings of the 2nd International Workshop on Semantics for Biodiversity (S4BioDiv)*. CEUR, 1933
- Michel F, Faron-Zucker C, Gandon F (2018) SPARQL Micro-Services: Lightweight Integration of Web APIs and Linked Data. *Proceedings of the Linked Data on the Web Workshop (LDOW2018)*.
- Parr C, Schulz K, Hammock J, Wilson N, Leary P, Rice J, Corrigan Jr R (2016) TraitBank: Practical semantics for organism attribute data. *Semantic Web* 7 (6): 577-588. <https://doi.org/10.3233/SW-150190>

Endnotes

- *1 19,300+ Web APIs are registered on ProgrammableWeb.com as of March 2018.